



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/082,354	02/26/2002	Brian Jacobsen	38627-170421	2756
26694	7590	09/29/2009	EXAMINER	
VENABLE LLP P.O. BOX 34385 WASHINGTON, DC 20043-9998				HICKS, MICHAEL J
ART UNIT		PAPER NUMBER		
		2165		
MAIL DATE		DELIVERY MODE		
09/29/2009		PAPER		

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	10/082,354	JACOBSEN ET AL.	
	<b>Examiner</b>	<b>Art Unit</b>	
	Michael J. Hicks	2165	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

1) Responsive to communication(s) filed on 13 July 2009.  
 2a) This action is **FINAL**.                    2b) This action is non-final.  
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

4) Claim(s) 1, 8, 24, 26-44, 46-48, 50-51, 54-59, and 62-66 is/are pending in the application.  
 4a) Of the above claim(s) 54-59 is/are withdrawn from consideration.  
 5) Claim(s) \_\_\_\_\_ is/are allowed.  
 6) Claim(s) 1, 8, 24, 26-44, 46-48, 50-51, and 62-66 is/are rejected.  
 7) Claim(s) \_\_\_\_\_ is/are objected to.  
 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

9) The specification is objected to by the Examiner.  
 10) The drawing(s) filed on \_\_\_\_\_ is/are: a) accepted or b) objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).  
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
 a) All    b) Some \* c) None of:  
 1. Certified copies of the priority documents have been received.  
 2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

1) Notice of References Cited (PTO-892)  
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)  
 3) Information Disclosure Statement(s) (PTO/SB/08)  
 Paper No(s)/Mail Date \_\_\_\_\_.  
 4) Interview Summary (PTO-413)  
 Paper No(s)/Mail Date \_\_\_\_\_.  
 5) Notice of Informal Patent Application  
 6) Other: \_\_\_\_\_.

**DETAILED ACTION**

1. Claims 1, 8, 24, 26-44, 46-48, 50-51, 54-59, and 62-66 Pending.

Claims 54-59 Withdrawn.

Claims 2-7, 9-23, 25, 45, 49, 52-53, and 60-61 Canceled.

***Continued Examination Under 37 CFR 1.114***

2. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114.

Applicant's submission filed on 7/13/2009 has been entered.

***Response to Arguments***

3. Applicant's arguments with respect to claims, 8, 24, 26-44, 46-48, 50-51, 54-59, and 62-66 have been considered but are moot in view of the new ground(s) of rejection.

***Claim Rejections - 35 USC § 102***

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. Claims 1, 8, 24, 26-34, 36, 38-44, 46-48, 50-51, and 62-66 rejected under 35 U.S.C. 102(b) as being anticipated by Diligenti et al. ("Focused Crawling Using Context Graphs", 26<sup>th</sup> International Conference on Very Large Databases, Pages 527-534, VLDB 2000 and referred to hereinafter as Diligenti).

As per Claims 1, 40 and 41, Diligenti discloses a computer-implemented method, system, and computer readable medium of implementing a specific search engine to compile and access subject-specific information, associated with a predefined particular subject, from a computer network, the method comprising the steps of:

traversing links between websites comprising one or more objects on the computer network, by said search engine, the objects respectively comprising at least one of: one or more web pages comprising the websites; and one or more components comprising any one or more of said web pages, the objects comprising at least one of: words, terms and expressions (See Page 5, Section 3.3 and Page 7, Column 1, Paragraph 2 which clearly disclose that a search engine may be used to initiate a crawl which traverses links between web pages, wherein the webpages are comprised of words terms and expressions.); filtering, by said search engine, subject specific contents of each site said object visited to determine a relevance of said subject specific content thereof to said predefined particular subject (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector

representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject.), and, wherein said filtering comprises: (a) decomposing said objects into one or more said components (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated.); (b) generating a lexicon comprising subject specific terminology deemed relevant to the predefined particular subject, the subject specific terminology comprising respective words, terms and expressions (See Page 4, Section 3.2 which clearly discloses that a classifier vector for a particular subject is computed which includes a vocabulary (e.g. lexicon) associated with that category comprising words terms and expressions.); (c) comparing said decomposed components of said objects to said subject specific terminology of the lexicon to determine whether each said object is a subject specific relevant object (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject.), wherein said comparing comprises: (i) assigning a weight to each of said words, terms and expressions comprising the subject specific terminology of the lexicon (See Page 4, Column 2 and Page 5, Column 1 which clearly disclose that each element of the vocabulary are assigned a weight, at least in that a number of matching words from the vocabulary is determined in the classification calculation (See Equation 3). As such it can be considered that each of the terms are equally weighted.); (ii) if a said word, term or expression comprising the

object matches a corresponding said word, term or expression comprising the subject specific terminology of the lexicon, adding a corresponding weight thereof to a cumulative total (See Page 5, Equation 3 which clearly discloses that the weight of each matching element is added to a cumulative total.); and (iii) determining any of said objects to be a subject specific relevant object if the cumulative total surpasses a predefined threshold value (See Page 5, Section 3.3 which clearly discloses that a confidence threshold is employed to determine relevant pages.); (d) based upon said comparing, determining all objects deemed to be subject specific relevant as objects to be saved (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.); presenting for an indexing operation at said search engine, each object determined to be site deemed subject specific relevant to said particular subject based upon said filtering (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.).

As per Claim 8, Diligenti discloses discarding all objects determined not to be subject specific relevant based upon said comparing (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed. Examiner notes that the objects deemed not relevant are discarded and not indexed.).

As per Claim 24, Diligenti discloses said filtering the occurs prior to said presenting (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3

further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject. Examiner notes that each of these steps occur before prior to results being returned or indexed as the express intent of the above steps is to determine what should be returned or indexed.).

As per Claim 26, Diligenti discloses replacing the lexicon with a lexicon corresponding to a different subject in order to present for said indexing operation create a different set of subject specific relevant objects subject-specific database (See Page 4, Section 3.2 which clearly discloses that multiple classifiers and therefore multiple vocabularies exists.).

As per Claim 27, Diligenti discloses the plurality of subject specific relevant objects are indexed and stored in a searchable database (See Page 7, Column 2, Paragraph 2 which clearly discloses that the method may be used as a search engine, and as such the results are saved in a searchable database.).

As per Claim 28, Diligenti discloses permitting a user to enter a query comprising user-preferred words, terms or expressions, wherein the steps of claim 1 are performed in response thereto (See Page 7, Column 2, Paragraph 2 which clearly discloses that the method may be used as a search engine.).

As per Claim 29, Diligenti discloses displaying information found in said step of searching (See Page 7, Column 2, Paragraph 2 which clearly discloses that the method may be used as a search engine.).

As per Claim 30, Diligenti discloses determining a site ranking for each website associated with information found in said searching step (See Page 7 Column 1, Paragraph 4-Column 2, Paragraph 1 which clearly discloses that the results are ranked and returned to the user.).

As per Claim 31, Diligenti discloses displaying the results of the user query using the site ranking of the information found in the searching step to determine an order in which the results are displayed (See Page 7 Column 1, Paragraph 4-Column 2, Paragraph 1 which clearly discloses that the results are ranked and returned to the user.).

As per Claim 32, Diligenti discloses displaying the results of the user query in a hierarchical format according to the site ranking (See Page 7 Column 1, Paragraph 4-Column 2, Paragraph 1 which clearly discloses that the results are ranked and returned to the user.).

As per Claim 33, Diligenti discloses for each said subject specific relevant stored in the database, assigning a word score to each word appearing on that object (See Page 4, Column 2 and Page 5, Column 1 which clearly disclose that each element of the vocabulary are assigned a weight, at least in that a number of matching words from the vocabulary is determined in the classification calculation (See Equation 3). As such it can be considered that each of the terms are equally weighted.).

As per Claim 34, Diligenti discloses determining all websites sites found in the database that contain links to the website site (See Page 6, Column 2, Paragraph 3 which clearly discloses that backlinks are determined and employed.); for each word on the websites site, assigning a word score for that word based at least in part on its presence on each website containing a link to the website site (See Page 4, Column 2 and Page 5, Column 1 which clearly disclose that each element of the vocabulary are assigned a weight, at least in that a number of matching words from the vocabulary is determined in the classification calculation (See Equation 3). As such it can be considered that each of the terms are equally weighted.).

As per Claim 36, Diligenti discloses determining all websites sites found in the database that contain links to the website See Page 6, Column 2, Paragraph 3 which clearly discloses that backlinks are determined and employed.); and assigning a word score to each word on the website site based at least in part on how many websites sites linking to the website site also contain the particular word (Examiner notes that the words are included in the vocabulary based upon a large number of sights including that words, and as such, the classifier inherently performs this function.).

As per Claim 38, Diligenti discloses entering a user query; using the user query to search the database (See Page 7 Column 1, Paragraph 4-Column 2, Paragraph 1 which clearly discloses that the results are ranked and returned to the user.); and computing a site ranking for each website site associated with information found in said searching step, the site ranking being computed based on said word scores (See

Page 7 Column 1, Paragraph 4-Column 2, Paragraph 1 which clearly discloses that the results are ranked and returned to the user.).

As per Claim 39, Diligenti discloses for each website site associated with information found in said searching step, summing the word scores for that website corresponding to words in the user query (See Page 5, Equation 3 which clearly discloses that the weight of each matching element is added to a cumulative total. Examiner notes that the user query may be used as the vocabulary.).

As per Claim 42, Diligenti discloses monitoring a depth for each said link, the depth being a reflection of relevance to said predefined particular subject (See Page 5, Column 1, Paragraph 1-2 which clearly discloses that the depth of the link is tracked and taken into account in the relevance judgment.).

As per Claim 43, Diligenti discloses for a given said object site being visited resulting from said link, setting a said depth of any links leading from said object that site to other objects to a depth of a link traversed to reach the given object (See Page 5, Column 1, Paragraph 1-2 which clearly discloses that the depth of the link is tracked and taken into account in the relevance judgment.); wherein said given object site is determined to be relevant to said predefined particular subject setting the depths of the links leading from said site to zero (See Page 5, Column 1, Paragraph 1-2 which clearly discloses that the depth of the link is tracked and taken into account in the relevance judgment.); and wherein said given object is determined not to be relevant to said predefined particular subject incrementing the depths of the links

leading from said object (See Page 5, Column 1, Paragraph 1-2 which clearly discloses that the depth of the link is tracked and taken into account in the relevance judgment.)..

As per Claim 44, Diligenti discloses comparing the incremented depths to a predetermined maximum depth value (See Page 5, Column 1, Paragraph 1-2 which clearly discloses that the depth of the link is tracked and taken into account in the relevance judgment and that a maximum depth exists.); wherein when the incremented depths exceed the predetermined maximum depth value, discarding the links leading from said given object (See Page 5, Column 1, Paragraph 1-2 which clearly discloses that the depth of the link is tracked and taken into account in the relevance judgment and that a maximum depth exists.); wherein when the incremented depths do not exceed the predetermined maximum depth value, traversing one of the links leading from said given objects (See Page 5, Column 1, Paragraph 1-2 which clearly discloses that the depth of the link is tracked and taken into account in the relevance judgment and that a maximum depth exists.).

As per Claim 46, Diligenti discloses a subject specific search engine system operable to compile and permit accessing of subject-specific information, associated with a predefined particular subject, from a computer network, the subject specific search engine system comprising: a host computer executing software stored upon a computer-readable storage medium, the software comprising: a subject specific smart crawler of said search engine traversing links between websites comprising one or more objects on the computer

network, the objects respectively comprising at least one of: one or more web pages comprising the websites; and one or components comprising any one or more of said web pages, the objects comprising at least one of: words, terms and expressions (See Page 5, Section 3.3 and Page 7, Column 1, Paragraph 2 which clearly disclose that a search engine may be used to initiate a crawl which traverses links between web pages, wherein the webpages are comprised of words terms and expressions.); said subject specific smart crawler performing filtering, a first filter of said search engine, to filter out sites, based on site contents, whose contents are irrelevant to said particular subject, and to permit only sites relevant to said particular subject to pass of subject specific content of each said object visited to determine a relevance of said subject specific content thereof to said predefined particular subject (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject. Examiner notes that sites deemed irrelevant (e.g. not meeting the minimum confidence threshold) are categorized as 'other' and not crawled further.), wherein said filtering comprises: (a) decomposing said objects into one or more said components (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated.); (b) generating a lexicon comprising subject specific terminology deemed relevant to the predefined particular subject, the subject specific terminology comprising respective words, terms and expressions (See Page 4, Section 3.2 which clearly

discloses that a classifier vector for a particular subject is computed which includes a vocabulary (e.g. lexicon) associated with that category comprising words terms and expressions.); (c) comparing said decomposed components of said objects to said subject specific terminology of the lexicon to determine whether each said object is a subject specific relevant object (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject.), wherein said comparing comprises: (i) assigning a weight to each of said words, terms and expressions comprising the subject specific terminology of the lexicon (See Page 4, Column 2 and Page 5, Column 1 which clearly disclose that each element of the vocabulary are assigned a weight, at least in that a number of matching words from the vocabulary is determined in the classification calculation (See Equation 3). As such it can be considered that each of the terms are equally weighted.); (ii) if a said word, term or expression comprising the object matches a corresponding said word, term or expression comprising the subject specific terminology of the lexicon, adding a corresponding weight thereof to a cumulative total (See Page 5, Equation 3 which clearly discloses that the weight of each matching element is added to accumulative total.); and (iii) determining any of said objects to be a subject specific relevant object if the cumulative total surpasses a predefined threshold value (See Page 5, Section 3.3 which clearly discloses that a confidence threshold is employed to determine relevant pages.); (d) based upon said comparing, determining all objects deemed to be subject specific relevant as

objects to be saved (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.); an indexer of said search engine indexing to index the relevant sites the plurality of said objects determined to be subject specific relevant to said particular subject based upon said filtering (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.); and a memory, connected to the host computer, for storing the plurality of said objects determined to be subject specific relevant (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.).

As per Claim 47, Diligenti discloses said filtering is performed by a first lexicon based filter (See Page 4, Section 3.2 which clearly discloses that a classifier vector for a particular subject is computed which includes a vocabulary (e.g. lexicon) associated with that category comprising words terms and expressions.).

As per Claim 48, Diligenti discloses the lexicon is stored on an interchangeable computer-readable storage medium (See Page 4, Section 3.2 which clearly discloses that a classifier vector for a particular subject is computed which includes a vocabulary (e.g. lexicon) associated with that category comprising words terms and expressions. Examiner notes that the classifiers, including the associated vocabularies are saved.).

As per Claim 50, Diligenti discloses the system further comprises a human-computer interface, and comprises: device for presenting said subject specific relevant objects received from the smart crawler to a human editor via

the human-computer interface (See Page 7, Column 2, Paragraph 2 which clearly discloses that the method may be used as a search engine, and as such the results are saved in a searchable database. Further, See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed. Examiner notes Page 2, Column 2, Paragraph 4 which clearly indicates that the process of focused crawling may be directed by a human user, as such Examiner asserts that the above disclosed step can be performed by a human editor via a human computer interface.); and device for receiving input from the human editor, entered via the human-computer interface, regarding whether to index and store said subject specific relevant objects in the memory (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed. Examiner notes Page 2, Column 2, Paragraph 4 which clearly indicates that the process of focused crawling may be directed by a human user, as such Examiner asserts that the above disclosed step can be performed by a human editor via a human computer interface.).

As per Claim 51, Diligenti discloses at least a second filter performing one or more operations of the first filter (See Page 5, Section 3.3 which clearly discloses that several classifiers may be utilized during the crawling.).

As per Claim 62, Diligenti discloses a computer-implemented method of implementing a subject specific search engine to compile and access subject specific information, associated with a predefined particular subject, from a computer network, the method comprising the steps of: traversing links between websites comprising one or more objects on the computer network, by said search engine, the objects respectively comprising at least one of: one or more

web pages comprising the websites; and one or components comprising any one or more of said web pages, the objects comprising at least one of: words, terms and expressions (See Page 5, Section 3.3 and Page 7, Column 1, Paragraph 2 which clearly disclose that a search engine may be used to initiate a crawl which traverses links between web pages, wherein the webpages are comprised of words terms and expressions.); filtering, by said search engine, subject specific content of each said object visited to determine relevance of said subject specific content thereof to said predefined particular subject (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject.), wherein said filtering comprises: (a) presenting one or more of said components of each of said objects to a human editor via a human computer interface (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject. Examiner notes that sites deemed irrelevant (e.g. not meeting the minimum confidence threshold) are categorized as 'other' and not crawled further. Examiner notes Page 2, Column 2, Paragraph 4 which clearly indicates that the process of focused crawling may be directed by a human user, as such Examiner asserts that the above disclosed step can be performed by a human editor via a human computer interface.); (b) permitting the human

editor to deem a said object to be a subject specific relevant object if the human editor determines any of said components comprising said object to be within said predefined particular subject (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject. Examiner notes that sites deemed irrelevant (e.g. not meeting the minimum confidence threshold) are categorized as 'other' and not crawled further. Examiner notes Page 2, Column 2, Paragraph 4 which clearly indicates that the process of focused crawling may be directed by a human user, as such Examiner asserts that the above disclosed step can be performed by a human editor via a human computer interface.); (c) permitting the human editor to deem a said object to not be a subject specific relevant object if the human editor determines any of said components comprising said object to not be within said predefined particular subject (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject. Examiner notes that sites deemed irrelevant (e.g. not meeting the minimum confidence threshold) are categorized as 'other' and not crawled further. Examiner notes Page 2, Column 2, Paragraph 4 which clearly indicates that the process of focused crawling may be directed by a human user, as such Examiner asserts that the above disclosed step can be performed by a human editor via a human computer interface.); and (d) based upon said (b) and (c), determining all objects deemed to be

subject specific relevant as objects to be saved (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.); presenting for an indexing operation at said search engine, each object determined to be subject specific relevant to said predefined particular subject based upon said filtering (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.).

As per Claim 63, Diligenti discloses a computer-implemented method of implementing a subject specific search engine to compile and access subject specific information, associated with a predefined particular subject, from a computer network, the method comprising the steps of: traversing links between websites comprising one or more objects on the computer network, by said search engine, the objects respectively comprising at least one of: one or more web pages comprising the websites; and one or components comprising any one or more of said web pages, the objects comprising at least one of: words, terms and expressions (See Page 5, Section 3.3 and Page 7, Column 1, Paragraph 2 which clearly disclose that a search engine may be used to initiate a crawl which traverses links between web pages, wherein the webpages are comprised of words terms and expressions.); filtering, by said search engine, subject specific content of each said object visited to determine relevance of said subject specific content thereof to said predefined particular subject (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2,

Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject.), wherein said filtering comprises (a) decomposing said objects into one or more said components (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated.); (b) generating a lexicon comprising subject specific terminology deemed relevant to the predefined particular subject, the subject specific terminology comprising respective words, terms and expressions (See Page 4, Section 3.2 which clearly discloses that a classifier vector for a particular subject is computed which includes a vocabulary (e.g. lexicon) associated with that category comprising words terms and expressions.); (c) comparing said decomposed components of said objects to said subject specific terminology of the lexicon to determine whether each said object is a subject specific relevant object, wherein a said object is deemed to be a subject specific relevant object if at least one component thereof matches a corresponding subject specific terminology of the lexicon (See Page 5, Section 3.3 which clearly discloses that for each retrieved and linked page, a reduced vector representation is calculated, Page 4, Column 2, Paragraph 3 clearly indicates that the reduced vector representation is comprised of components (e.g. words terms and expressions) from the object, and Page 5, Section 3.3 and Page 4, Column 2, Paragraph 3 further make it clear that this vector is compared against a classifier vector to determine relevancy for the object to a predefined particular subject.); (d) based upon said comparing, determining all objects deemed to be subject specific relevant as objects to be saved (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.); presenting for an indexing operation at said search engine, each object determined to be subject specific relevant to said predefined particular subject

based upon said filtering (See Page 7, Column 1, Paragraph 2 which clearly discloses that the objects deemed relevant are saved and indexed.).

As per Claim 64, Diligenti discloses indexing the totality of objects determined to be subject specific relevant to yield a subcategory of objects (See Page 3, Section 3 and Page 6, Column 2, Paragraphs 4 which clearly disclose that multiple classifiers exists which are defined for multiple categories.).

As per Claim 65, Diligenti discloses the objects are websites, the computer network comprises the Internet, and the subcategory of objects comprises a portion of the Internet (Internet') (See Page 7, Column 2, Paragraph 2 which clearly discloses that the method may be used as a search engine, and as such the results are saved in a searchable database.).

As per Claim 66, Diligenti discloses performing a searching operation upon the Internet' (See Page 7, Column 2, Paragraph 2 which clearly discloses that the method may be used as a search engine, and as such the results are saved in a searchable database.).

### ***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been

obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 35 and 37 are rejected under 35 U.S.C. 103(a) as being unpatentable over Diligenti in view of Menczer et al. (as cited in the Final Office Action dated 12/11/2008).

As to claims 35, and 37, Diligenti fails to disclose the step of assigning a word score for that word further comprises the step of increasing the word score for each site containing a link to the site when the word appears in close proximity to the link.

Menczer et al. discloses the step of assigning a word score for that word further comprises the step of increasing the word score for each site containing a link to the site when the word appears in close proximity to the link (See Menczer et al. page 244, column 1, paragraph 1, and page 246, column 1, paragraph 2, wherein it is read to mean "link appearance" calculations).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to further modify the invention of Diligenti as modified by the teachings of Menczer et al. to include the step of assigning a word score for that word further comprises the step of increasing the word score for each site containing a link to the site when the word appears in close proximity to the link because ranking and scoring content is a well known method in organizing and providing relevant query results.

***Points of Contact***

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael J. Hicks whose telephone number is (571) 272-2670. The examiner can normally be reached on Monday - Friday 9:00a - 5:30p.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Neveen Abel-Jalil can be reached at (571)272-4074. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Michael J Hicks  
Art Unit 2165  
Phone: (571) 272-2670  
Fax: (571) 273-2670

/Neveen Abel-Jalil/

Supervisory Patent Examiner, Art Unit 2165